# 春节特别放送2｜用自顶至底的思路解决数据统计问题

2021-02-13 尹会生

Python自动化办公实战课 进入课程 >

**讲述：尹会生**
时长 05:05 大小 4.67M

你好，我是尹会生。今天是大年初二，先祝你春节快乐。

上节课我策划了一个项目作业，你完成得怎么样了呢？如果你对第一模块的内容掌握了60%，其实这道题你就能完成得差不多了，或者说至少会有一个思路。那么今天这节课，我也来给你提供一些我的解题思路和解题方法。

我们再来回顾下两道题目：

1. 统计出全年每种水果的总销售额。
2. 统计出每月销售数量排名 Top 3 的水果。

因为这两道题目都是需要在统计数据的前提下进行的，一个是统计之后进行求和，一个是统计之后进行排序。所以，我们优先要做的就是统计全部文件所有 sheet 里的单元格数据。

**以第一题为例，可以按照自顶至底的思路拆分问题。**

什么是自顶至底呢？意思是分解问题的时候，先关注总体，再解决具体问题。我给你举个例子你会更容易理解。这是一个你熟悉的问题：怎样把大象装进冰箱？

你会马上想到，把大象装进冰箱分成三个步骤。你看，你不会关注冰箱是什么品牌，大象是什么品种，冰箱的容积如何，大象的体积有多大等等。这样思考的好处是你可以先对高层次的问题进行定义、设计、编码，再细节的问题，再其中的子任务或下一层次去解决。这样逐层去进行设计和编码能够让你的程序有更好的可读性。

那统计销售总数，我也可以把问题拆分成三个步骤。

第一步，拆解 Excel，通过第二讲的循环功能，取得每日销售数据。

第二步，读取每天里的每次销售金额数据。

第三步，将数据按照自己的算法（这里就是简单的累加）进行汇总。

接下来我带你具体看下每个步骤需要做什么。

首先，我们来取出每日销售数据。

你可以先从总销售额开始入手，思考如何拆解问题,。总销售额可以拆分成按月销售额的累加，每月的销售数据可以存放到单独的文件中。月销售额可以拆分成日销售额的累加，日销售额放在每个 sheet 中。

如果用 python 来实现总销售额的统计，实现思路应该为：通过 pathlib 库和 for 循环遍历文件，读取每个 Excel 文件，再通过 xlrd 库读取每个 Excel 里的每个 sheet，每个 sheet 里记录了每一天的销售数据, 这样可以把每天的销售数据读取出来。

接着，我们来读取每一次销售数据。

由于 Excel 中除了销售金额，还有日期等其他噪声数据，我们只需要读取“水果名称”“销售金额”这两列必须数据，就能实现统计销售总额的功能了。因此通过这两个数据所在的列取出“水果名称”“销售金额”这两列。

用 python 实现的话, 就是：使用 for 循环遍历每一行的水果，再采用“字典”数据类型临时存放到 Python 的一种数据结构中。为什么要使用“字典”呢？在第三讲我给你讲解过，“字典”适合记录有映射关系的一对或多对数据。

最后，将数据汇总在一起。

根据需求，销售的金额只需要记录汇总的金额就可以实现最终要求，也就是统计全年销售额的任务了。所以这一步具体做法就是：把字典中可以对每种水果的每次销售金额累加存放，就可以得到当日每种水果销售总金额了。

第二道题，那如何来实现 Top3 功能呢?

这两道问题的处理思路非常相似，都是需要处理每个月的销售额，但是第一个问题，统计全年销售额，需要将 12 个月的销售数据进行累加。第二个问题，需要对每个月的销售数据进行从大到小的排序，前三位的就是“水果销售 Top3”了。

你的思路和想法是什么呢？可以在留言区分享一下，我们一起交流、讨论。在 2 月 16 日，也就是大年初五这一天，我会把完整答案给你分享出来，供你参考。

提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。 页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 春节特别放送1 | 实体水果店转线上销售的数据统计问题

下一篇 春节特别放送3 | 揭晓项目作业的答案

## 精选留言（1）

写留言

**巩春雨**
2021-02-15

这个数据不是很多，可不可以合并成一张表，然后进行筛选求和呢